

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	The smoking habit of a close friend or family member – how deep is the impact? A cross-sectional study
<b>AUTHORS</b>	Saari, Antti; Kentala, Jukka; Mattila, Kari

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Vanessa Johnston Senior Research Fellow Menzies School of Health Research I have no competing interests.
<b>REVIEW RETURNED</b>	04-Jun-2013

<b>THE STUDY</b>	<p>One overall comment I have is that in parts, the standard of written English is not at a high enough standard for an academic publication. I can only imagine that it must be challenging to write an academic publication in a language that is not your first language, but the style in sections is confusing to the reader and makes interpretation of the results and discussion, in particular, difficult. I would recommend that this paper is edited by a native English speaker if possible.</p> <p>I believe some of the conclusions are too far reaching for the results presented in this study, particularly in relation to the conclusions drawn about the influence of family and friends on initiation, duration and intensity which are all based on univariate analyses.</p> <p>There needs to be more detail in the statistical analysis section, including on adjustment for potential confounders and how missing data and loss to follow-up was addressed.</p> <p>In parts (I have highlighted where in my review) the authors need to draw more strongly on the literature to back up their statements.</p> <p>I cannot judge the external validity of this study based on the information presented.</p>
<b>RESULTS &amp; CONCLUSIONS</b>	<p>I believe some of the conclusions are too far reaching for the results presented in this study, particularly in relation to the conclusions drawn about the influence of family and friends on initiation, duration and intensity which are all based on univariate analyses.</p>
<b>GENERAL COMMENTS</b>	<p>This paper's main aim is to ascertain the risk of being an adult smoker associated with parental smoking, sibling and best friend smoking at age 13 years in a Finnish cohort.</p> <p>This is a large cohort study and potentially has some interesting insights to offer but I believe the paper needs some major revisions in order to add some robust new evidence to the evidence base in</p>

	<p>this area.</p> <p>One overall comment I have is that in parts, the standard of written English is not at a high enough standard for an academic publication. I can only imagine that it must be challenging to write an academic publication in a language that is not your first language, but the style in sections is confusing to the reader and makes interpretation of the results and discussion, in particular, difficult. I would recommend that this paper is edited by a native English speaker if possible.</p> <p><b>Introduction</b></p> <p>Overall the introduction is quite well written.</p> <p>The 3rd paragraph describing the evidence on which social influences are most important in determining youth uptake and smoking progression is not as clear cut as the authors present. My reading of the literature is that the effect of family smoking is particularly relevant for younger children, whereas peer group behaviours are more important in influencing smoking during teenage years. More recent longitudinal research suggests parental influences are important for initiation and escalation of smoking. Peer behaviour too, has been found to affect initiation, progression and trajectories.</p> <p>If the evidence was so clear about which social influences are most relevant and at which time points, then what is the rationale for this study? Other than perhaps to test the established evidence in the Finnish context. I think the rationale for this study could be made more strongly.</p> <p>The 4th paragraph sets out the social theory behind social influence and smoking behaviour. This again needs some more detail. In perceived influence theories, is smoking among peers perceived as the norm? How does smoking behaviour of peers directly influence smoking – through peer pressure or some other/additional mechanism?</p> <p>I note that in the discussion the authors state they were only interested in whether smokers <i>perceived</i> they had smokers among their family and friends – presumably, referencing the perceived influence theories they introduce in the introduction. This needs to be described in the methods.</p> <p>In the 6<sup>th</sup> paragraph the authors introduce the cohort study conducted in Finland. At first the relevance of this was unclear to me. There needs to be a stronger link between this study and the aim of the current paper in the introduction. The details about the cohort study should be moved to the methods. What does “very homogenous school conditions” mean?</p> <p>The stated aim does not capture other results that are presented in the paper: the association between smoking attitudes at age 12 and</p>
--	--

	<p>smoking and the association of family/friends smoking and age of initiation/duration/intensity of smoking.</p> <p><b>Methods</b></p> <p>I would suggest describing the details of the questions that were asked regarding smoking in the methods section (currently repeated in intro and methods). Was there any data about parental income/education/occupation, as this might be a confounding factor? Some background about why a birth cohort had a dental check-up at age 12 would be interesting. Is this a universal health check-up in Finland?</p> <p>Calculation of duration of smoking for smokers and ex-smokers was quite crude as it did not take into account any quit periods – this should be acknowledged in limitations.</p> <p>Number of cigarettes was used as a measure of intensity of consumption. What about roll-your-own cigarettes made from loose tobacco? How was this accounted for?</p> <p>I am uncertain why participants were asked in 2008 about the smoking behaviour of their parents when this was asked in 1992? What data on parental smoking was used for the different analyses?</p> <p>Was the author's measure of self-rated health a standardised measure?</p> <p>"In this study we combined two different settings" is a confusing statement. I think what is meant is "In this study we combined two different research designs."</p> <p>I don't believe the second "setting" was retrospective; it was a cross-sectional analysis only. Can the authors please confirm or clarify this.</p> <p>There needs to be more detail in the statistical analysis section, including on adjustment for potential confounders and how missing data and loss to follow-up was addressed.</p> <p><b>Results</b></p> <p>As the authors note there was a large loss to follow up. It would be useful to see a comparison between those retained and those lost (by gender, smoking status at age 13, smoking attitudes) to ascertain how different these groups are. I see from Table 1 that the % of women at follow-up was greater than at baseline. This should also be acknowledged in discussion when discussing loss to follow up.</p> <p>Table 1 should include the denominators in the first row and should consider adding the denominators for smoking behavior. I would suggest adding median or mean age in this table as well.</p>
--	---

	<p>The writing style in the following sentence is confusing and needs some editing:</p> <p>“Smokers also differed statistically significantly from non-smokers with regard to their self-perceived health; more than one out of four smokers had not very good self-perceived health while only one of ten non-smokers had other than very good self-perceived health (Table 2). There were also statistically significantly more females among non-smokers (Table 2)” (p.9)</p> <p>Table 2 – heading should include the year 2008 somewhere for clarity</p> <p>“Prospective setting” should read “prospective study.”</p> <p>These following few sentences are confusing. Do the authors mean “Twenty-four percent of smokers in 2008 reported having a smoking parent in 1992, compared to 11.7% of non-smokers”?</p> <p>“There were 24.1% (n=99) smokers in 2008 among those who reported having a smoking parent or smoking parents in the 1992 questionnaire, while only 11.7% (n=68) of those with the opposite response smoked (<math>p&lt;0.001</math>). Females did not differ statistically significantly from males according to the results on this question: OR for being smoker was 2.0 (1.3–3.3) for females and 2.6 (1.5–4.7) for males with smoking parent(s).” (p.10-11)</p> <p>Why was an OR used and not relative risk?</p> <p>“Cross-sectional setting” should read “cross-sectional study.”</p> <p>The following sentences I believe should be moved to the prospective section.</p> <p>“Having a smoker as a best friend in school was connected with more than five times greater likelihood of being a smoker in adulthood in females, but not in males. Smoking of mother, brother or sister when the subject was of school age increased the likelihood of being a smoker more than twofold among males, but not among females. The smoking behavior of father when the subject was of school age did not have a connection with smoking behavior in adulthood. (Table 3)”(p.11-12)</p> <p>Table 3 – I don’t suggest tabulating univariate and multivariate analyses together in the same table. Consider summarising the univariate data in text and only tabulating the multivariate analysis.</p> <p>Should include self-perceived health and marital status in multivariate table.</p> <p>Table 4 is confusing. Are the comparisons between smokers and ex-smokers or for smoker/ex-smoker participants who answered yes and no to each of the criteria (e.g. best friend smoking)</p> <p>Analyses on age of initiation, duration and intensity are all univariate</p>
--	---

	<p>and I don't believe you can draw strong conclusions from them, as they are not adjusted for potential confounders.</p> <p><b>Discussion</b></p> <p>I think the discussion is the weakest section of this paper. I believe some of the conclusions are too far reaching for the results presented in this study, particularly in relation to the conclusions drawn about the influence of family and friends on initiation, duration and intensity which are all based on univariate analyses.</p> <p>The limitations too are not adequately addressed. The response rate is low and this will bias results as demonstrated by the fact that more women were in the follow-up sample. Additionally, the majority of this sample had higher education and had very good self-rated health. What about recall and reporting bias that are potential sources of bias in studies that rely on self-reported questionnaires.</p> <p>The authors introduce new information on p.17 in the discussion – that half of the cohort received brief tobacco interventions during school. This needed to come in earlier (in methods, when describing the original cohort study). It is unclear what the authors mean by “recent evidence of cessation interventions is in line with our findings.” (p.17 para 3)</p> <p>The interpretation of the findings in the context of existing literature also needs to be strengthened. The first paragraph has no referencing attached to multiple statements about the role of friends in influencing smoking behaviour.</p> <p>Para 2, p.16 – references a Finnish study that concurs with the results of this study. What about international literature?</p> <p>While the concluding paragraph describes how smoking prevention needs to be a comprehensive, multi-pronged approach, it does not address strongly enough potential points of intervention based on what this study has found – that is, the influence of friends and the friendship group on smoking behaviour.</p> <p>As to the external validity of this study, I am not convinced. As stated, the majority of this sample had higher education and very good self-rated health. How do these variables, as well as smoking prevalence among the cohort in 2008 compare with the wider Finnish population? Is it representative?</p>
--	---

<b>REVIEWER</b>	<p>Dr. J.M. Vink Dept. Biological Psychology VU University, Amsterdam The Netherlands No competing interests.</p>
<b>REVIEW RETURNED</b>	14-Jul-2013

<p><b>THE STUDY</b></p>	<p>Introduction</p> <p>The overview of papers describing the relation between smoking initiation and smoking family members of friends is not complete. The authors do not mention the possibility that family members show similar smoking behavior due to shared genes. Papers exploring the association between smoking family members and the uptake of smoking, taking genetic factors into account, should be included in the introduction (for example: Twin Research 2003, 6(3), pp. 209-217 and/or twin and heritability studies)</p> <p>The authors mention that smoking rate among young adult females remains a significant issue , but do not mention what the prevalence of smoking is in this group (in general prevalence of smoking seems to be lower in females than males according to the previous sentence in the introduction).</p> <p>Method:</p> <p>Duration of smoking was calculated for smokers by subtracting age at initiation from 29 -&gt; but what if someone started smoking at 16, quit at 18 but started again at 28? Than he/she only smoked for 3 years, but according to this calculation it will be 13 years... Same question for the calculation for ex-smokers.</p> <p>Only t-tests or chi-square tests were used. Did the authors consider a regression analyses including all available variables?</p>
-------------------------	---

<b>RESULTS &amp; CONCLUSIONS</b>	<p><b>Results</b></p> <p>Smoking rates are lower (14.4% in females and 20.4% in males) in the study sample compared to the general population (1 in 4 males and 1 in 5 females smokes according to the introduction). Is this due to response bias?</p> <p>Authors conclude that some associations (for example influence of smoking mother on smoking in males) are significant (for example <math>p=0.04</math>) but did they consider to correct for multiple testing?</p> <p>To predict the influence of smoking family members and best friend on smoking behavior in adulthood, it might be (more) interesting to select never-smokers at the age of 13 instead of including the total group.</p> <p>It might be interesting to explore the influence of smoking mother, father, brother, sister on smoking behavior separately for males and females because of possible sex differences. See also previous literature on this topic.</p> <p>Question about Table 4: what happened if someone did not have a brother or a sister? Was this person not included in the analyses? Were all analyses done separately for mother, father, brother, sister etcetera? Did the authors consider to do a regression analyses including all variables at the same time to see whether significant associations remained significant when corrected for the other variables? And if yes, how much variance was explained by this model?</p> <p><b>Discussion</b></p> <p>Why do the authors think their results can be generalized to industrialized populations? Sample size is rather small, results are inconclusive?</p> <p>In my opinion, the suggestions that are described in the last part of the discussion (about smoking prevention and the role of doctors and others) cannot be concluded from the results described in this paper.</p> <p>There is much more literature on this topic (influence of smoking family and friends), which should be included in the introduction/discussion. Authors should make clear what their results add to the existing literature.</p>
----------------------------------	--

## VERSION 1 – AUTHOR RESPONSE

First the comments by Dr. Vanessa Johnston:

*1. One overall comment I have is that in parts, the standard of written English is not at a high enough standard for an academic publication. I can only imagine that it must be challenging to write an academic publication in a language that is not your first language, but the style in sections is confusing to the reader and makes interpretation of the results and discussion, in particular, difficult. I would recommend that this paper is edited by a native English speaker if possible.*

It is true that English language is not our first language, but the text was in fact checked by native English speaker, likewise this revised version.

*2. The 3rd paragraph describing the evidence on which social influences are most important in determining youth uptake and smoking progression is not as clear cut as the authors present. My reading of the literature is that the effect of family smoking is particularly relevant for younger children, whereas peer group behaviours are more important in influencing smoking during teenage years. More recent longitudinal research suggests parental influences are important for initiation and escalation of smoking. Peer behaviour too, has been found to affect initiation, progression and trajectories. If the evidence was so clear about which social influences are most relevant and at which time points, then what is the rationale for this study? Other than perhaps to test the established evidence in the Finnish context. I think the rationale for this study could be made more strongly.*

Comparing the different social effects was the point of this study. We now explain it more thoroughly in the last paragraph of the Introduction.

*3. The 4th paragraph sets out the social theory behind social influence and smoking behaviour. This again needs some more detail. In perceived influence theories, is smoking among peers perceived as the norm? How does smoking behaviour of peers directly influence smoking – through peer pressure or some other/additional mechanism? I note that in the discussion the authors state they were only interested in whether smokers perceived they had smokers among their family and friends – presumably, referencing the perceived influence theories they introduce in the introduction. This needs to be described in the methods.*

Our study was based on perceived influence theory. This is now described in the third paragraph of the Methods section.

*4. In the 6th paragraph the authors introduce the cohort study conducted in Finland. At first the relevance of this was unclear to me. There needs to be a stronger link between this study and the aim of the current paper in the introduction. The details about the cohort study should be moved to the methods. What does “very homogenous school conditions” mean?*

This information has now been moved to the first paragraph of the Methods section. The word “homogenous” has been replaced with “similar”. This refers to the common rural orientation, demographics, and economic conditions in these schools.

*5. The stated aim does not capture other results that are presented in the paper: the association between smoking attitudes at age 12 and smoking and the association of family/friends smoking and age of initiation/duration/intensity of smoking.*



We decided to omit the results about age at initiation and duration of smoking, likewise the results about the heaviness of smoking. Thus these are not included in the study aims. Smoking attitudes are now only used for analysis of nonresponse and thus are not mentioned in the study aims.

#### *Methods*

*6. I would suggest describing the details of the questions that were asked regarding smoking in the methods section (currently repeated in intro and methods).*

These details now remain only in the Methods section.

*7. Was there any data about parental income/education/occupation, as this might be a confounding factor?*

Unfortunately these confounders were not measured during data collection. This is a limitation of our study and is thus mentioned in the Discussion, new “Limitations” section.

*8. Some background about why a birth cohort had a dental check-up at age 12 would be interesting. Is this a universal health check-up in Finland?*

In Finland all children had dental check-ups annually during the 1990's. This is not mentioned in the text since it does not affect the results in any way.

*9. Calculation of duration of smoking for smokers and ex-smokers was quite crude as it did not take into account any quit periods – this should be acknowledged in limitations.*

We no longer mention the duration of smoking in our manuscript. We agree that our calculation method was rather imprecise.

*10. Number of cigarettes was used as a measure of intensity of consumption. What about roll-your-own cigarettes made from loose tobacco? How was this accounted for?*

Heaviness of smoking is no longer mentioned in the manuscript. The respondents were, however, asked to estimate their tobacco consumption in cigarettes if they used pipe or loose tobacco.

*11. I am uncertain why participants were asked in 2008 about the smoking behaviour of their parents when this was asked in 1992? What data on parental smoking was used for the different analyses?*

The questions differed from each other. In 1992 they were asked if one or both of their parents smoked and in 2008 if their mother smoked when they were of school age with a similar question about their father. The question asked in 1992 is now used for analysis of nonresponse (see additions to the Methods section). In the Results we use solely the responses to the 2008 follow-up.

*12. Was the author's measure of self-rated health a standardised measure?*

Self-rated health has been used in many cross-sectional and cohort studies we know of. For example, the HeSSup study with its numerous publications uses the same method for grading self-perceived health. For example see: Suominen S, Koskenvuo K, Sillanmäki L, Vahtera J, Korkeila K, Mattila K, Virtanen P, Sumanen M, Rautava P, Koskenvuo M. Non-response in a nationwide follow-up postal survey in Finland: A register-based mortality analysis of respondents and non-respondents of the Health and Social Support (HeSSup) Study. *BMJ open*; 2012;2:e000657

*13. "In this study we combined two different settings" is a confusing statement. I think what is meant is "In this study we combined two different research designs."*

Now we use only the cross-sectional design and this sentence has been removed from the paper.

*14. I don't believe the second "setting" was retrospective; it was a cross-sectional analysis only. Can the authors please confirm or clarify this.*

We now renamed the setting cross-sectional instead of retrospective. This change has been made to the title as well.

*15. There needs to be more detail in the statistical analysis section, including on adjustment for potential confounders and how missing data and loss to follow-up was addressed.*

We now provide more detail in the Statistical analysis section.

## *Results*

*16. As the authors note there was a large loss to follow up. It would be useful to see a comparison between those retained and those lost (by gender, smoking status at age 13, smoking attitudes) to ascertain how different these groups are.*

We have done the advised analysis and the results are presented as Analysis of non-response (in the Methods section).

*17. I see from Table 1 that the % of women at follow-up was greater than at baseline. This should also be acknowledged in discussion when discussing loss to follow up.*

We now acknowledge the selection bias of females in Limitations (Discussion).

18. *Table 1 should include the denominators in the first row and should consider adding the denominators for smoking behavior. I would suggest adding median or mean age in this table as well.*

We consider the table (renamed Table 2) quite self-explanatory as it is. We added mean ages to the row headings.

19. *The writing style in the following sentence is confusing and needs some editing:*

*“Smokers also differed statistically significantly from non-smokers with regard to their self-perceived health; more than one out of four smokers had not very good self-perceived health while only one of ten non-smokers had other than very good self-perceived health (Table 2). There were also statistically significantly more females among non-smokers (Table 2)” (p.9)*

Language help was given to us by a native English speaker, who corrected the sentence for greater clarity.

20. *Table 2 – heading should include the year 2008 somewhere for clarity “Prospective setting” should read “prospective study.”*

The prospective analysis section has now been removed from the paper and this heading no longer exists.

21. *These following few sentences are confusing. Do the authors mean “Twenty-four percent of smokers in 2008 reported having a smoking parent in 1992, compared to 11.7% of non-smokers”? “There were 24.1% (n=99) smokers in 2008 among those who reported having a smoking parent or smoking parents in the 1992 questionnaire, while only 11.7% (n=68) of those with the opposite response smoked ( $p<0.001$ ). Females did not differ statistically significantly from males according to the results on this question: OR for being smoker was 2.0 (1.3–3.3) for females and 2.6 (1.5–4.7) for males with smoking parent(s).” (p.10-11)*

Yes, we did, but now we no longer use these results in the paper.

22. *Why was an OR used and not relative risk?*

OR comes from the logistic regression analysed using SPSS for Windows v. 20.0. We think both OR and RR represent the same phenomenon and are both acceptable, but only use one in the same paper.

23. *“Cross-sectional setting” should read “cross-sectional study.”*

This has been corrected.

*24. The following sentences I believe should be moved to the prospective section. "Having a smoker as a best friend in school was connected with more than five times greater likelihood of being a smoker in adulthood in females, but not in males. Smoking of mother, brother or sister when the subject was of school age increased the likelihood of being a smoker more than twofold among males, but not among females. The smoking behavior of father when the subject was of school age did not have a connection with smoking behavior in adulthood. (Table 3)"(p.11-12)*

We now present all results under one heading and there no longer exists a prospective study – section. We hope this improves the readability of our paper.

*25. Table 3 – I don't suggest tabulating univariate and multivariate analyses together in the same table. Consider summarising the univariate data in text and only tabulating the multivariate analysis. Should include self-perceived health and marital status in multivariate table.*

We now have only the multivariate analysis in the table, univariate results are summarised in the Results. We also added education, self-perceived health and marital status to the multivariate table.

*26. Table 4 is confusing. Are the comparisons between smokers and ex-smokers or for smoker/ex-smoker participants who answered yes and no to each of the criteria (e.g. best friend smoking) Analyses on age of initiation, duration and intensity are all univariate and I don't believe you can draw strong conclusions from them, as they are not adjusted for potential confounders.*

The table has been removed from the paper since we decided to concentrate on the primary aim of the study and improve the clarity of our findings.

### *Discussion*

*27. I think the discussion is the weakest section of this paper. I believe some of the conclusions are too far reaching for the results presented in this study, particularly in relation to the conclusions drawn about the influence of family and friends on initiation, duration and intensity which are all based on univariate analyses.*

Now we no longer discuss the results that were based on the univariate analyses since these results have also been removed from the Results section.

*28. The limitations too are not adequately addressed. The response rate is low and this will bias results as demonstrated by the fact that more women were in the follow-up sample. Additionally, the*

*majority of this sample had higher education and had very good self-rated health. What about recall and reporting bias that are potential sources of bias in studies that rely on self-reported questionnaires.*

We now mention these limitations in the Limitations section. However, the majority of Finns born in 1970's actually have higher education so the representativeness of our study population can be considered satisfactory.

*29. The authors introduce new information on p.17 in the discussion – that half of the cohort received brief tobacco interventions during school. This needed to come in earlier (in methods, when describing the original cohort study).*

We now mention the brief tobacco interventions in the Methods section.

*30. It is unclear what the authors mean by “recent evidence of cessation interventions is in line with our findings.”(p.17 para 3)*

This paragraph has been omitted.

*31. The interpretation of the findings in the context of existing literature also needs to be strengthened. The first paragraph has no referencing attached to multiple statements about the role of friends in influencing smoking behaviour.*

We now include some references to support our insights, which have also been reworded for greater clarity.

*32. Para 2, p.16 – references a Finnish study that concurs with the results of this study. What about international literature?*

We think the setting of the referenced study was so similar to the one we used that its concurring results greatly enhanced the reliability of our results. The matter has – of course – been reported in many other journals from which we could have picked a reference.

*33. While the concluding paragraph describes how smoking prevention needs to be a comprehensive, multi-pronged approach, it does not address strongly enough potential points of intervention based on what this study has found – that is, the influence of friends and the friendship group on smoking behaviour.*

This matter has now been more thoroughly addressed at the end of the Discussion section.

*34. As to the external validity of this study, I am not convinced. As stated, the majority of this sample had higher education and very good self-rated health. How do these variables, as well as smoking prevalence among the cohort in 2008 compare with the wider Finnish population? Is it representative?*

As mentioned earlier, we consider the population representative. The majority of Finns take higher education.

Second, the comments by Dr. Jacqueline M. Vink:

*35. The overview of papers describing the relation between smoking initiation and smoking family members of friends is not complete. The authors do not mention the possibility that family members show similar smoking behavior due to shared genes. Papers exploring the association between smoking family members and the uptake of smoking, taking genetic factors into account, should be included in the introduction (for example: Twin Research 2003, 6(3), pp. 209-217 and/or twin and heritability studies).*

We made an addition with a reference to the Introduction about the aspect of shared genes.

*36. The authors mention that smoking rate among young adult females remains a significant issue , but do not mention what the prevalence of smoking is in this group (in general prevalence of smoking seems to be lower in females than males according to the previous sentence in the introduction).*

We now explain the trend more clearly in the Introduction.

*Method:*

*37. Duration of smoking was calculated for smokers by subtracting age at initiation from 29 -> but what if someone started smoking at 16, quit at 18 but started again at 28? Than he/she only smoked for 3 years, but according to this calculation it will be 13 years... Same question for the calculation for ex-smokers.*

We decided to omit the analysis considering the duration of smoking and/or age at initiation from the paper to clarify our message.

*38. Only t-tests or chi-square tests were used. Did the authors consider a regression analysis including all available variables?*

We did indeed conduct a binary logistic regression for the essential variables (see for example Table 3 in the previous and Table 4 in this revised version) and now we tabulate only the multivariate analysis including those found to be potential confounders (statistically different measurements among smokers and non-smokers).

### *Results*

*39. Smoking rates are lower (14.4% in females and 20.4% in males) in the study sample compared to the general population (1 in 4 males and 1 in 5 females smokes according to the introduction). Is this due to response bias?*

Probably. This matter has now been added to the Limitations section in the Discussion.

*40. Authors conclude that some associations (for example influence of smoking mother on smoking in males) are significant (for example  $p=0.04$ ) but did they consider to correct for multiple testing?*

We did these multivariate analyses but the associations were non-significant after adjustment for confounders. We chose to omit some of the Results (including this) to clarify the point of this paper. Thus this conclusion has also been removed.

*41. To predict the influence of smoking family members and best friend on smoking behavior in adulthood, it might be (more) interesting to select never-smokers at the age of 13 instead of including the total group. It might be interesting to explore the influence of smoking mother, father, brother, sister on smoking behavior separately for males and females because of possible sex differences. See also previous literature on this topic.*

We think the family influences begin to affect individuals' health behavior (including attitudes towards smoking) much earlier and thus such a never-smokers study should be initiated at a very young age, even before school. We agree that this kind of approach would be interesting.

*42. Question about Table 4: what happened if someone did not have a brother or a sister? Was this person not included in the analyses?*

Those without a brother or a sister were excluded from the analysis on the family member in question since they could not respond anything on the question. We made an addition about this to the Methods.

*43. Were all analyses done separately for mother, father, brother, sister etcetera? Did the authors consider to do a regression analyses including all variables at the same time to see whether significant associations remained significant when corrected for the other variables? And if yes, how much variance was explained by this model?*

We did this analysis (see Table 3 in the previous version and Table 4 in the revised version), but variance was not included since we present 95% confidence intervals which also demonstrates the variance.

#### *Discussion*

*44. Why do the authors think their results can be generalized to industrialized populations? Sample size is rather small, results are inconclusive?*

Our population is representative on Finns and thus we now agree that generalisation could only be done to Finnish populations and comparable populations. This is mentioned in the Discussion.

*45. In my opinion, the suggestions that are described in the last part of the discussion (about smoking prevention and the role of doctors and others) cannot be concluded from the results described in this paper.*

*There is much more literature on this topic (influence of smoking family and friends), which should be included in the introduction/discussion.*

Conclusions have now been extended to include the key points of this particular study.

*46. Authors should make clear what their results add to the existing literature.*

This has been explained more thoroughly in "What this paper adds". It has earlier been known that family influences and peer influences are important in smoking behavior. In our setting it is possible to compare the effects of genetic and epigenetic influences for smoking in adulthood.